

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/768,080

DATE: 07/30/2001

TIME: 15:09:22

Input Set : A:\GC527C3-seqlist.txt

Output Set: N:\CRF3\07302001\I768080.raw

3 <110> APPLICANT: Estell, David
 4 Harding, Fiona
 6 <120> TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND
 7 METHODS OF MAKING AND USING THE SAME
 9 <130> FILE REFERENCE: GC527C3
 11 <140> CURRENT APPLICATION NUMBER: US 09/768,080
 12 <141> CURRENT FILING DATE: 2001-01-23
 14 <150> PRIOR APPLICATION NUMBER: US 09/677,822
 15 <151> PRIOR FILING DATE: 2000-10-02
 17 <150> PRIOR APPLICATION NUMBER: US 09/500,135
 18 <151> PRIOR FILING DATE: 2000-02-08
 20 <150> PRIOR APPLICATION NUMBER: US 09/060,872
 21 <151> PRIOR FILING DATE: 1998-04-15
 23 <160> NUMBER OF SEQ ID NOS: 240
 25 <170> SOFTWARE: PatentIn Ver. 2.1
 27 <210> SEQ ID NO: 1
 28 <211> LENGTH: 1495
 29 <212> TYPE: DNA
 30 <213> ORGANISM: Bacillus amyloliquefaciens
 32 <220> FEATURE:
 33 <221> NAME/KEY: mat_peptide
 34 <222> LOCATION: (417)..(1495)
 36 <220> FEATURE:
 37 <221> NAME/KEY: CDS
 38 <222> LOCATION: (96)..(1244)
 40 <220> FEATURE:
 41 <221> NAME/KEY: misc_feature
 42 <222> LOCATION: (582)..(584)
 43 <223> OTHER INFORMATION: The nnn at positions 582 through 584 which in a
 44 preferred embodiment (aat) is to code for
 45 asparagine, but which may also code for proline.
 47 <220> FEATURE:
 48 <221> NAME/KEY: misc_feature
 49 <222> LOCATION: (585)..(587)
 50 <223> OTHER INFORMATION: The nnn at positions 585 through 587 which in a
 51 preferred embodiment (cct) is to code for proline,
 52 but which may also code for asparagine.
 54 <220> FEATURE:
 55 <221> NAME/KEY: misc_feature
 56 <222> LOCATION: (597)..(599)
 57 <223> OTHER INFORMATION: The nnn at positions 597 to 599 which in a
 58 preferred embodiment (aac) is to code for
 59 asparagine, but which may also code for aspartic acid.
 61 <220> FEATURE:
 62 <221> NAME/KEY: misc_feature
 63 <222> LOCATION: (678)..(680)
 64 <223> OTHER INFORMATION: The nnn at positions 678 through 680 which in a

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/768,080

DATE: 07/30/2001
TIME: 15:09:22

Input Set : A:\GC527C3-seqlist.txt
Output Set: N:\CRF3\07302001\I768080.raw

125	aaa tat att gtc ggg ttt	aaa cag aca atg	agc acg atg	agc gcc gct	257				
126	Lys Tyr Ile Val Gly Phe	Lys Gln Thr Met	Ser Thr Met	Ser Ala Ala					
127	-65	-60	-55						
129	aag aag aaa gat gtc att	tct gaa aaa ggc ggg	aaa gtg caa aag	caa aag	305				
130	Lys Lys Lys Asp Val Ile	Ser Glu Lys Gly	Gly Lys Val Gln	Lys Gln					
131	-50	-45	-40						
133	ttc aaa tat gta gac gca	gtc tca gct aca	tta aac gaa	aaa gct gta	353				
134	Phe Lys Tyr Val Asp Ala	Ala Ser Ala Thr	Leu Asn Glu	Lys Ala Val					
135	-35	-30	-25						
137	aaa gaa ttg aaa aaa gac	ccg agc gtc gct	tac gtt gaa gaa	gat cac	401				
138	Lys Glu Leu Lys Lys Asp	Pro Ser Val Ala	Tyr Val Glu	Glu Asp His					
139	-20	-15	-10						
141	gta gca cat gcg tac	gcg cag tcc	gtg cct tac	ggc gta tca	caa att	449			
142	Val Ala His Ala Tyr	Ala Gln Ser Val	Pro Tyr Gly	Val Ser Gln	Ile				
143	-5	-1 1	5	10					
145	aaa gcc cct gct ctg	cac tct caa	ggc tac act	gga tca aat	gtt aaa	497			
146	Lys Ala Pro Ala Leu	His Ser Gln Gly	Tyr Thr Gly	Ser Asn Val	Lys				
147	15	20	25						
149	gta gcg gtt atc	gac agc ggt	atc gat	tct tct	cat cct gat	tta aag			
150	Val Ala Val Ile Asp	Ser Gly Ile Asp	Ser Ser His	Pro Asp Leu	Lys				
151	30	35	40						
W--> 153	gta gca ggc gga	gcc agc atg	gtt cct tct	gaa aca nnm nnm	ttc caa	593			
W--> 154	Val Ala Gly Gly	Ala Ser Met Val	Pro Ser Glu	Thr Xaa Xaa	Phe Gln				
155	45	50	55						
W--> 157	gac nnm aac tct	cac gga act	cac gtt	gcc ggc aca	gtt gcg gct	641			
W--> 158	Asp Xaa Asn Ser His	Gly Thr His Val	Ala Gly	Thr Val Ala	Ala Leu				
159	60	65	70	75					
W--> 161	aat aac tca atc	ggt gta tta	ggc gtt	gct cca	agc nnm nnm	ctt tac	689		
W--> 162	Asn Asn Ser Ile	Gly Val Leu	Gly Val Ala	Pro Ser Xaa	Xaa Leu	Tyr			
163	80	85	90						
W--> 165	gct gta aaa gtt	ctc ggt nnm	nnm ggt	tcc ggc	caa tac	agc tgg atc	737		
W--> 166	Ala Val Lys Val	Leu Gly Xaa	Xaa Gly	Ser Gly	Gln Tyr	Ser Trp	Ile		
167	95	100	105						
169	att aac gga atc	gag tgg	gct atc	gca aac	aat atg	gac gtt	att aac	785	
170	Ile Asn Gly Ile	Glu Trp Ala	Ile Ala Asn	Asn Met	Asp Val	Ile Asn			
171	110	115	120						
173	atg agc ctc	ggc gga	cct tct	ggt tct	gct gct	tta aaa	gct gca gtt	833	
174	Met Ser Leu Gly	Gly Pro Ser	Gly Ser Ala	Ala Leu	Lys Ala	Ala Val			
175	125	130	135						
177	gat aaa gcc	gtt gca tcc	ggc gtc	gtc gtt	gct gca	gcc ggt	aaa aac	881	
178	Asp Lys Ala Val	Ala Ser Gly	Val Val Val	Ala Ala	Ala Gly	Asn			
179	140	145	150	155					
W--> 181	gaa ggc nnm	nnm ggc	agc tca	agc aca	gtg ggc	tac cct	ggt aaa tac	929	
W--> 182	Glu Gly Xaa Xaa	Gly Ser Ser	Ser Thr Val	Gly Tyr Pro	Gly Lys	Tyr			
183	160	165	170						
185	cct tct gtc	att gca	gta ggc	gct gtt	gac agc	agc aac	caa aga	gca	977
186	Pro Ser Val Ile	Ala Val Gly	Ala Val Asp	Ser Ser Asn	Gln Arg	Ala			
187	175	180	185						
189	tct ttc tca	agc gta	gga cct	gag ctt	gat gtc	atg gca	cct ggc	gta	1025

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/768,080

DATE: 07/30/2001

TIME: 15:09:22

Input Set : A:\GC527C3-seqlist.txt
 Output Set: N:\CRF3\07302001\I768080.raw

65 preferred embodiment (gca) is to code for
 66 alanine, but which may also code for serine.
 68 <220> FEATURE:
 69 <221> NAME/KEY: misc_feature
 70 <222> LOCATION: (681)..(683)
 71 <223> OTHER INFORMATION: The nnn at positions 681 through 683 which in a
 72 preferred embodiment (tca) is to code for serine,
 73 but which may also code for alanine.
 75 <220> FEATURE:
 76 <221> NAME/KEY: misc_feature
 77 <222> LOCATION: (708)..(710)
 78 <223> OTHER INFORMATION: The nnn at positions 708 through 710 which in a
 79 preferred embodiment (gct) is to code for
 80 alanine, but which may also code for aspartic acid.
 82 <220> FEATURE:
 83 <221> NAME/KEY: misc_feature
 84 <222> LOCATION: (711)..(713)
 85 <223> OTHER INFORMATION: The nnn at positions 711 through 713 which in a
 86 preferred embodiment (gac) is to code for
 87 aspartic acid, but which may also code for alanine.
 89 <220> FEATURE:
 90 <221> NAME/KEY: misc_feature
 91 <222> LOCATION: (888)..(890)
 92 <223> OTHER INFORMATION: The nnn at positions 888 through 890 which in a
 93 preferred embodiment (act) is to code for
 94 threonine, but which may also code for serine.
 96 <220> FEATURE:
 97 <221> NAME/KEY: misc_feature
 98 <222> LOCATION: (891)..(893)
 99 <223> OTHER INFORMATION: The nnn at positions 891 through 893 which in a
 100 preferred embodiment (tcc) is to code for
 101 serine, but which may also code for threonine.
 103 <220> FEATURE:
 104 <221> NAME/KEY: misc_feature
 105 <222> LOCATION: (1167)..(1169)
 106 <223> OTHER INFORMATION: The nnn at positions 1167 through 1169 which in
 107 a preferred embodiment (gaa) is to code for
 108 glutamic acid, but which may also code for glutamine.
 110 <400> SEQUENCE: 1
 111 ggtctactaa aatattatttc catactatac aattaataca cagaataatc tgtctattgg 60
 113 ttattctgca aatgaaaaaaaa aggagaggat aaaga atg aga ggc aaa aaa gta 113
 114 Met Arg Gly Lys Lys Val
 115 -105
 117 tgg atc agt ttg ctg ttt gct tta gcg tta atc ttt acg atg gcg ttc 161
 118 Trp Ile Ser Leu Leu Phe Ala Leu Ala Leu Ile Phe Thr Met Ala Phe
 119 -100 -95 -90
 121 ggc agc aca tcc tct gcc cag gcg gca ggg aaa tca aac ggg gaa aag 209
 122 Gly Ser Thr Ser Ser Ala Gln Ala Ala Gly Lys Ser Asn Gly Glu Lys
 123 -85 -80 -75 -70

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/768,080

DATE: 07/30/2001

TIME: 15:09:22

Input Set : A:\GC527C3-seqlist.txt

Output Set: N:\CRF3\07302001\I768080.raw

190 Ser Phe Ser Ser Val Gly Pro Glu Leu Asp Val Met Ala Pro Gly Val
 191 190 195 200
 193 tct atc caa aac acg ctt cct gga aac aaa tac ggg gcg tac aac ggt 1073
 194 Ser Ile Gln Ser Thr Leu Pro Gly Asn Lys Tyr Gly Ala Tyr Asn Gly
 195 205 210 215
 197 acg tca atg gca tct ccg ccc gtt gcc gga gcg gct gct ttg att ctt 1121
 198 Thr Ser Met Ala Ser Pro His Val Ala Gly Ala Ala Leu Ile Leu
 199 220 225 230 235
 W--> 201 tct aag cac ccg aac tgg aca aac act caa gtc cgc agc agt tta ~~nnn~~ 1169
 W--> 202 Ser Lys His Pro Asn Trp Thr Asn Thr Gln Val Arg Ser Ser Leu Xaa
 203 240 245 250
 205 aac acc act aca aaa ctt ggt gat tct ttc tac tat gga aaa ggg ctg 1217
 206 Asn Thr Thr Lys Leu Gly Asp Ser Phe Tyr Tyr Gly Lys Gly Leu
 207 255 260 265
 209 atc aac gta cag gct gca gct cag taa aacataaaaa acggcccttg 1264
 210 Ile Asn Val Gln Ala Ala Gln
 211 270 275
 213 qccccggggg tttttttatt ttttttccctt cggcatgttca atccgctcca taatcqacgg 1324
 215 atggctccctt ctgaaaaattt taacgagaaaa cggccgggtt acccggttca gtcccgtaac 1384
 217 qcccaagtcc ttgaaacgttca caatccggccgc ttcccggtt ccgggtcaatccgtaa 1444
 219 acgggtccggcg gcttttccctt gataccggga gacggcattc gtaatccggat c 1495
 222 <211> SEQ ID NO: 2
 223 <211> LENGTH: 382
 224 <212> TYPE: PRT
 225 <213> ORGANISM: Bacillus amyloliquefaciens
 227 <220> FEATURE:
 228 <221> NAME/KEY: VARIANT
 229 <222> LOCATION: (163)...(163)
 230 <223> OTHER INFORMATION: Xaa = Asn or Pro
 232 <220> FEATURE:
 233 <221> NAME/KEY: VARIANT
 234 <222> LOCATION: (164)...(164)
 235 <223> OTHER INFORMATION: Xaa = Pro or Asn
 237 <220> FEATURE:
 238 <221> NAME/KEY: VARIANT
 239 <222> LOCATION: (167)...(167)
 240 <223> OTHER INFORMATION: Xaa = Asn or Asp
 242 <220> FEATURE:
 243 <221> NAME/KEY: VARIANT
 244 <222> LOCATION: (195)...(195)
 245 <223> OTHER INFORMATION: Xaa = Ala or Ser
 247 <220> FEATURE:
 248 <221> NAME/KEY: VARIANT
 249 <222> LOCATION: (196)...(196)
 250 <223> OTHER INFORMATION: Xaa = Ser or Ala
 252 <220> FEATURE:
 253 <221> NAME/KEY: VARIANT
 254 <222> LOCATION: (205)...(205)
 255 <223> OTHER INFORMATION: Xaa = Ala or Asp

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/768,080

DATE: 07/30/2001
TIME: 15:09:22

Input Set : A:\GC527C3-seqlist.txt
Output Set: N:\CRF3\07302001\I768080.raw

257 <220> FEATURE:
258 <221> NAME/KEY: VARIANT
259 <222> LOCATION: (206)...(206)
260 <223> OTHER INFORMATION: Xaa = Asp or Ala
262 <220> FEATURE:
263 <221> NAME/KEY: VARIANT
264 <222> LOCATION: (265)...(265)
265 <223> OTHER INFORMATION: Xaa = Thr or Ser
267 <220> FEATURE:
268 <221> NAME/KEY: VARIANT
269 <222> LOCATION: (266)...(266)
270 <223> OTHER INFORMATION: Xaa = Ser or Thr
272 <220> FEATURE:
273 <221> NAME/KEY: VARIANT
274 <222> LOCATION: (358)...(358)
275 <223> OTHER INFORMATION: Xaa = Gln or Glu
277 <400> SEQUENCE: 2
278 Met Arg Gly Lys Val Trp Ile Ser Leu Leu Phe Ala Leu Ala Leu
279 1 5 10 15
280 Ile Phe Thr Met Ala Phe Gly Ser Thr Ser Ser Ala Gln Ala Ala Gly
281 20 25 30
282 Lys Ser Asn Gly Glu Lys Lys Tyr Ile Val Gly Phe Lys Gln Thr Met
283 35 40 45
284 Ser Thr Met Ser Ala Ala Lys Lys Asp Val Ile Ser Glu Lys Gly
285 50 55 60
286 Gly Lys Val Gln Lys Gln Phe Lys Tyr Val Asp Ala Ala Ser Ala Thr
287 65 70 75 80
288 Leu Asn Glu Lys Ala Val Lys Glu Leu Lys Lys Asp Pro Ser Val Ala
289 85 90 95
290 Tyr Val Glu Glu Asp His Val Ala His Ala Tyr Ala Gln Ser Val Pro
291 100 105 110
292 Tyr Gly Val Ser Gln Ile Lys Ala Pro Ala Leu His Ser Gln Gly Tyr
293 115 120 125
294 Thr Gly Ser Asn Val Lys Val Ala Val Ile Asp Ser Gly Ile Asp Ser
295 130 135 140
296 Ser His Pro Asp Leu Lys Val Ala Gly Gly Ala Ser Met Val Pro Ser
297 145 150 155 160
W--> 298 Glu Thr Xaa Xaa Phe Gln Asp Xaa Asn Ser His Gly Thr His Val Ala
299 165 170 175
300 Gly Thr Val Ala Ala Leu Asn Asn Ser Ile Gly Val Leu Gly Val Ala
301 180 185 190
W--> 302 Pro Ser Xaa Xaa Leu Tyr Ala Val Lys Val Leu Gly Xaa Xaa Gly Ser
303 195 200 205
304 Gly Gln Tyr Ser Trp Ile Ile Asn Gly Ile Glu Trp Ala Ile Ala Asn
305 210 215 220
306 Asn Met Asp Val Ile Asn Met Ser Leu Gly Gly Pro Ser Gly Ser Ala
307 225 230 235 240
308 Ala Leu Lys Ala Ala Val Asp Lys Ala Val Ala Ser Gly Val Val Val
309 245 250 255

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/768,080

DATE: 07/30/2001

TIME: 15:09:23

Input Set : A:\GC527C3-seqlist.txt

Output Set: N:\CRF3\07302001\I768080.raw

L:153 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:154 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:157 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:158 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:161 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:162 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:165 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:166 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:181 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:182 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:201 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:202 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:298 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:302 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:310 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:322 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2